
Characterization of human DNA sequences synthesized at the onset of S-phase

Carla Tribioli, Giuseppe Biamonti, Mauro Giacca, Marco Colonna, Silvano Riva and Arturo Falaschi*

Istituto di Genetica Biochimica ed Evoluzionistica del CNR, Via Abbiategrasso, 207, I-27100 Pavia, Italy

Received September 25, 1987; Revised and Accepted November 25, 1987

ABSTRACT

We developed a method of enrichment for DNA replicated at the onset of S-phase in synchronized human HL60 cells. About 200 such sequences were cloned. The analysis of this selected DNA sample showed that: 1) the cloned DNA fragments derive from a limited number (750-1500) of replicons; 2) there is no extensive homology between different DNA fragments; 3) they are not significantly enriched in highly repeated sequences; 4) they are enriched in snap-back (Cot = 0) DNA. The sequence of the longest fragment revealed the presence of numerous signals collected in a few hundred nucleotides: 1) homology with the origin of replication of human Papovaviruses usually associated with potential stem-loop structures; 2) binding sites for known transcription factors and for another nuclear factor; 3) potential binding sites for the chromosome "scaffold".

INTRODUCTION

The replication of eukaryotic genomes occurs by the activation of multiple tandemly organized replication units (replicons) each containing a bidirectional origin. Somatic animal cell chromosomes contain several thousand replicons (one per 50 to 100 kb on the average). Replicons that are coordinately activated are usually arranged in clusters which are distributed among and within the chromosomes (1-4). Only a fraction of them replicates at any given time during the S-phase and the temporal order of replication is maintained from one cell cycle to the next but may vary with development and differentiation (4-6). These and other findings suggest that replicons might be correlated with or possibly correspond to functional genetic domains (7-9). The key constitutive element of a replicon is the origin which, by analogy with the prokaryotic equivalent, is thought to be a specific nucleotide sequence where DNA replication starts and where the

Nucleic Acids Research

control mechanisms act. Origin activation probably requires a critical and transient sequence of preceding events, a proper dosage of regulatory factors and a suitable chromatin structure (5,10). In addition many experiments point to a correlation between initiation of replication and transcription (11-14).

Experiments performed in different laboratories indicate that cloned animal cell origins are inactive when re-introduced into the same cells by transfection (15). Thus, in the absence of a functional assay such as that used in yeast to isolate autonomously replicating sequences (ARS), the only practicable approaches are at the moment those based on the enrichment for replication origins by physical methods. One such approach is based on the synchronization of cells at the G1/S border and the purification and cloning of the DNA synthesized at the onset of S-phase (16). Using this last approach we have in effect cloned such DNA from human cells. The longest fragment was sequenced and found to contain homologies with the known replication origins of human Papovavirus, putative regions of attachment to the chromosome "scaffold" and recognition sites for transcription factors.

MATERIALS AND METHODS

Materials

Restriction enzymes, DNA polymerase I Klenow fragment, T4 kinase and proteinase K were purchased from Boehringer Corp. Mannheim, F.R.G. and used as recommended by the manufacturer. DNase I and S1 nuclease were obtained from Worthington Corp. Waltham, Mass. U.S.A.; Mung bean nuclease was obtained from Promega Biotec, Madison, Wis. U.S.A. BrdUrd and dNTPs were purchased from Boehringer Corp. γ - 32 P-ATP and α - 32 P dCTP (both at 3000 Ci/mmol), 3 H-BrdUrd, (25 Ci/mmol), 3 H-dThd (25 Ci/mmol) and 14 C-dThd (60 mCi/mmol) were all obtained from Amersham Corp., U.K.

Labeling of DNA with α - 32 P dCTP

Radioactive probes for hybridization experiments were prepared either by conventional nick-translation (17) or utilizing the multiprimer DNA labeling technique (18) as specified by the manufacturer (Amersham). By this last method a specific activity of $1-2 \times 10^9$ cpm/ μ g DNA was obtained.

Synchronization of cells at the G1/S border

Human HL60 lymphoblastoid cells were synchronized by a modification of the procedure of Pedrali-Noy et al (19). Cells were grown at a density of 5×10^5 cells/ml in RPMI 1640 medium (GIBCO Lab. Grand Island, N.Y.) containing: 10 $\mu\text{g/ml}$ gentamicin, 10% fetal calf serum, 2mM L-glutamine. When required, cells were uniformly labeled with ^{14}C -dThd (60 mCi/mmole) at a final concentration of 5 $\mu\text{Ci/ml}$, 0.1 μM . Synchronization was started by adding an equal volume of prewarmed medium containing 2 $\mu\text{g/ml}$ aphidicolin. After 24 hours of incubation (37° C) the drug was removed by centrifugating at 2000 xg for 10 min and washing once with medium without aphidicolin at room temperature. Cells were then resuspended in medium without aphidicolin at a density of 2×10^5 cells/ml and incubated for 12 hours at 37°C. Aphidicolin was then added at 5 $\mu\text{g/ml}$ and incubation continued for 12 hours. Cells were finally collected by centrifugation, washed twice with fresh medium without the drug and immediately used in labeling experiments.

Labelling of DNA synthesized at the onset of S-phase

Cells synchronized at the G1/S border were resuspended at 10^7 cells/ml in fresh RPMI 1640 medium supplemented as described above, (except fetal calf serum) containing 1 $\mu\text{g/ml}$ aphidicolin and either 1 μM ^3H -dThd (25 Ci/mmol) or 100 μM ^3H -BrdUrd (1.5 Ci/mmole) as specified in the text. After a time ranging from 10 to 130 min (see text) at 37°C, cells were killed with 0.02% Na azide, collected by centrifugation (2000 xg, 10 min. at 4°C) and washed twice with cold Phosphate Buffer Saline (PBS). Radioactivity was measured on 10 μl of culture after precipitation with cold 10% TCA. All manipulation of Br-labeled DNA were performed under dim red lighth.

Extraction of high molecular weight DNA from synchronized cells after labeling with ^3H -dThd

High molecular weight DNA was extracted according to Maniatis et al (17) from nuclei purified as described by Challberg and Kelly (20).

Extraction of DNA labeled with ^3H -BrdUrd from synchronized cells

Nuclei, purified as described by Challberg and Kelly (20) were resuspended at 5×10^6 /ml in: 10 mM Tris-HCl pH8.0, 10 mM EDTA, 0.5% SDS, 100 $\mu\text{g/ml}$ proteinase K. After incubating 3 hours at 37°C the solution was

rendered 200 mM NaCl and forced 10 times through a siringe needle (25 gauge). The DNA so treated was directly used for fractionation in CsCl density gradient centrifugations as described in the legend to Fig. 3.

Treatment with Mung bean nuclease, blunt-end ligation and transformation of *E. coli*.

25 ng of (HL) DNA or 50 ng of (HH) DNA prepared as described above were incubated with Mung bean nuclease (1.15 U) in the presence of 1.5 μ g of dephosphorylated vector DNA (pAT153 cut at NruI site) in 50 mM Na acetate, pH5.0, 0.1 mM ZnCl₂. Incubation was at 0°C for 30 min. Digestion was stopped by rendering the solution 10 mM EDTA pH8.0 and 10 μ g/ml yeast tRNA. Half of the treated DNA was then phenol extracted, ethanol precipitated and ligated with T4 ligase under conditions that optimize blunt-end ligation (17). The ligation mixtures were then used to transform *E. coli* strain HB101 according to Hanahan (21) and all transformants were collected.

Isolation of DNA synthesized in different S-phase intervals

5×10^8 synchronized HL60 cells were collected by centrifugation, washed with 150 ml fresh medium and resuspended in 1 litre of prewarmed RPMI 1640 medium containing 10 μ g/ml gentamicin, 10% dialyzed fetal calf serum, 2 mM L-glutamine. 133 ml of culture were immediately centrifugated 10 min at 3000 xg, washed with 50 ml cold PBS and the cellular pellet was stored at -80°C. This sample represented cells in the G1 phase. The rest of the culture was divided into six 133 ml aliquots that were further incubated at 37°C. The S-phase was divided into six intervals: 0-60 min, 60-120 min, 120-210 min, 210-300 min, 300-390 min, 390-450 min. Each time-interval sample was labeled with 0.1 μ M ³H-dThd (1 μ Ci/ml), in the presence of 100 μ M cold BrdUrd. At the end of each time-interval, cells were killed with Na azide, collected by centrifugation, washed with PBS and stored at -80°C. DNA was purified as already described for ³H-BrdUrd labeled DNA. (LL) DNA and (HL) DNA were separated by CsCl density gradient centrifugation as already described. (LL) DNA was detected by OD 260 nm readings while (HL) DNA was detected by radioactivity measurements on the fractions. (LL) and (HL) DNA containing fractions were pooled dialyzed and concentrated.

Renaturation kinetics analysis

The method described by Pedrali-Noy et al (22) was followed. Briefly human HL60 cells uniformly labeled with ^{14}C -dThd (60 mCi/mmole) at a final concentration of 5 $\mu\text{Ci/ml}$, 0.1 μM , were synchronized by the method previously described. DNA synthesis was then allowed for 7 min in the absence of aphidicolin and ^{14}C -dThd and in presence of ^3H -dThd (25 Ci/mmole) at a final concentration of 50 $\mu\text{Ci/ml}$. High molecular weight DNA was then extracted according to Maniatis et al (17) and resuspended in 10 mM Tris-HCl pH 8.0, 1 mM EDTA at a final concentration of 800 $\mu\text{g/ml}$. 300 μl of DNA (240 μg) were boiled for 15 min in order to produce small fragments (average 500 nucleotides) in 0.3 M NaOH, neutralized by adding 16.5 μl of 6 M HCl, and then divided into 25 μl aliquots in 25 μl capillaries. The capillaries were incubated for different times at 60° C. After 0, 1, 5, 10, 15, 20, min and 2, 4, 6, 11, 20 and 22 hrs the capillaries were chilled in ice, opened and the content was digested for 3 hrs at 37°C with 15 U of S1 nuclease in 60 mM Na succinate pH 4.9, 1.5 mM Zn SO₄, native calf thymus DNA 40 $\mu\text{g/ml}$ (final volume 100 μl). The fraction of S1 resistant DNA (^3H and ^{14}C) was determined as TCA precipitable material.

Southern blot, dot blot, slot blot hybridization and DNA sequencing

Southern blot hybridizations were performed according to Maniatis et al. (17). Dot blot and slot blot experiments were according to Mariani and Schimke (23) with a Schleicher and Schuell, Dassel F.R.G., apparatus. Hybridization conditions were as follows: 6x SSC, 10x Denhardt's solution, 0.1% SDS, 100 $\mu\text{g/ml}$ yeast tRNA, 68°C for 48 hrs; two washing at 68° C in 6 x SSC 0.1% SDS, one washing in 0.2 x SSC, 0.1% SDS at 68°C for 1 hr. DNA sequencing was according to Maxam and Gilbert (24).

Band shift experiments and DNase I footprinting analysis

Binding reactions were carried out by incubating 1 ng of the end-labeled fragment (10.000 cpm) with 5 μg of nuclear extract from HeLa or HL60 cells, prepared as described by Dignam et al. (25), and 3 μg of poly [d(I-C)]: poly [d(I-C)] (Boehringer) in 20 mM Hepes pH 7.3, 50 mM NaCl, 5 mM MgCl₂, 2 mM DTT, 0.2 mM EDTA, 4 mM spermidine, 5% glycerol (final volume: 20 μl); nuclear extracts were added last (26). After 30 min incubation at room temperature,

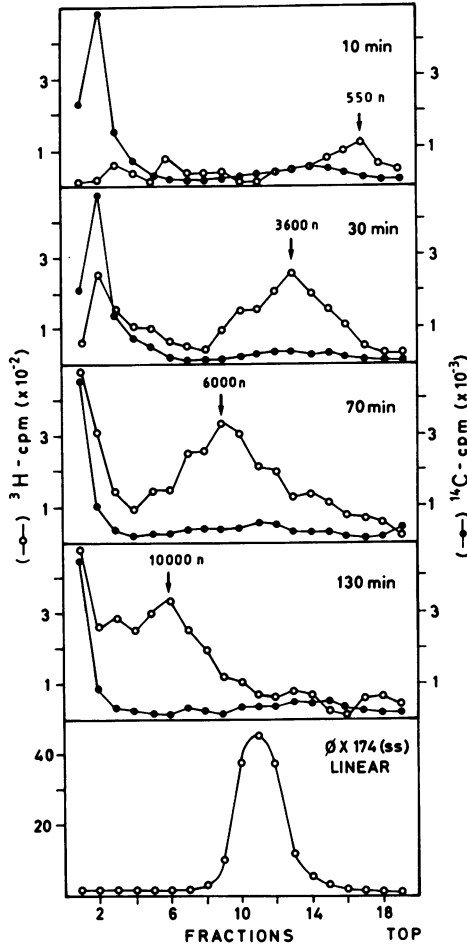


Fig. 1: Sucrose gradient centrifugation of newly synthesized DNA. High Mr DNA labeled with ^{14}C - and ^3H -dTdp and purified as described in Materials and Methods from synchronized cells was resuspended in 100 μl TE buffer and denatured by adding NaOH (0.5 M final). After 10 min in ice, DNA was dialyzed against TE buffer pH8.0 at 4°C for 3 hours and then loaded onto 5 ml 5-30% sucrose gradients in 10 mM Tris-HCl, pH8.0, 1 mM EDTA, 0.3 M NaCl. Gradients were run 3 hours at 45 000 rpm in a Beckman SW50 rotor. The \bar{s} value was determined by comparison with that of linear single-stranded $\phi\text{x}174$ DNA (5.37 Kb, 28 \bar{S}). The molecular weight of DNA was calculated from the corresponding \bar{s} value according to Studier (27).

TABLE I

| Estimate of the number of activated replicons | | | | |
|---|------------------------------------|--|--|--------------------|
| Labeling time (min) | ^3H -cpm per 10^5 cells | ^3H -DNA (pg/cell) ^(a) | Size of fragments ^(b) (nucleotides) | Fragments per cell |
| 10 | 44 | 6.0×10^{-4} | 550 | 2000 |
| 30 | 164 | 2.1×10^{-3} | 3600 | 1000 |
| 70 | 417 | 5.4×10^{-3} | 6000 | 1600 |
| 130 | 648 | 8.4×10^{-3} | 10000 | 1500 |

(a) Calculated on the basis of a measured specific activity of 7.7×10^5 cpm/ μg .
 (b) From the experiment in Fig. 1.

samples were loaded onto a 5% polyacrylamide gel and analyzed according to Carthew *et al* (26).

DNase I footprinting experiments were performed essentially as described by Carthew *et al.* (26). 10 ng (10^5 cpm) of the 195 bp Bgl I - Ava I fragment of plasmid pB48 (see text) labeled at the Ava I end were incubated with 50 μg of HeLa nuclear extract under the binding conditions described above (final volume: 100 μl). After 30 min incubation at room temperature, freshly diluted DNase I was added at a final concentration of 5 ng/ μl , and the digestion was allowed to proceed 60 sec at room temperature. The products of the reaction were analyzed as previously described (26).

RESULTS

Synchronization, kinetics of DNA synthesis

The synchronization procedure for HL60 cells and labeling of newly synthesized DNA is described in Materials and Methods. The rate of increase in length of newly synthesized DNA was measured by neutral sucrose gradient centrifugation after alkaline denaturation. As Fig. 1 shows ^3H -labeled DNA sediments separately from the bulk of ^{14}C -DNA (unreplicated) and

progressively increases in length from 500 to 10 000 bp between 10 and 130 min after release from block. This indicates that in our conditions the new DNA synthesis was mostly replicative with some repair synthesis only at the latest times, if at all and that the rate of fork advancement was reduced to about 2 % of the normal rate. A comparison of the overall rate of DNA synthesis with that of chain elongation (see Table 1) leads to an interesting conclusion: the increase of DNA synthesized per cell can for the most part be accounted by the increase in size of a fixed number of molecules (about 1500).

Renaturation kinetics analysis

The results of this experiment are shown in Fig. 2 where the renaturation kinetics of parental (^{14}C) and newly synthesized DNA (^3H) are compared. As the figure shows, newly synthesized DNA is significantly enriched in snap-back ($\text{Cot} = 0$) structures as compared to parental DNA (5.6 % Vs. 3.5 %). On the other hand the other rapidly renatured DNA (highly repeated sequences) seems to be equally represented in the two DNAs. This result, obtained with a very stringent technique (resistance to S1 nuclease digestion) is in agreement with those obtained by other authors with hydroxyapatite chromatography on a different system (28). Thus it appears that the enrichment for snap-back DNA is a distinctive feature of the DNA replicated after synchronization with aphidicolin.

Fractionation of newly synthesized DNA labeled with ^3H -BrdUrd

A preparative experiment was performed, in which newly synthesized DNA was density labeled with ^3H -BrdUrd. The heavy DNA (^3H) was separated from the bulk of unreplicated DNA (^{14}C) by CsCl density gradient centrifugations. (see Fig. 3). Most of ^3H -labeled DNA banded at heavy-light (HL) density, however a fraction of heavy-heavy (HH) DNA was also present. Both (HL) and (HH) DNA fractions were pooled and further purified by successive CsCl gradient centrifugations. The presence of (HH) DNA could be due either to reinitiation from the same origins (29) or to extrusion of nascent DNA molecules in the double stranded form due to branch migration in sheared DNA (30).

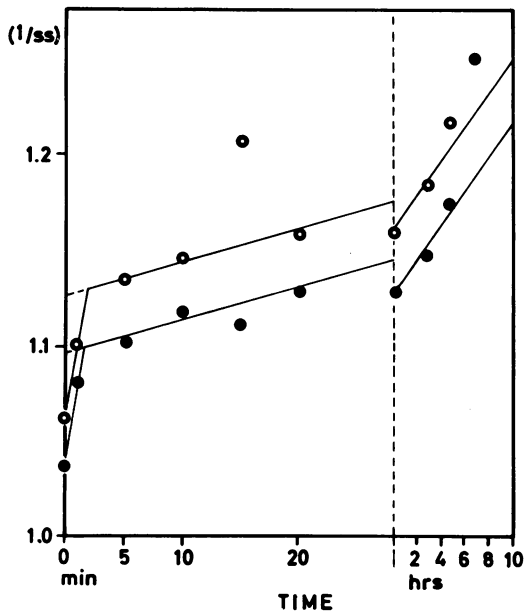


Fig. 2: Renaturation kinetics analysis.

The procedure described in Materials and Methods was followed. The reciprocal of the fraction of S1 nuclease susceptible DNA (1/ss) is plotted against renaturation time for both newly synthesized DNA (○) and parental DNA (●).

Cloning of nascent DNA fragments into pAT153

The purified (HL) and (HF) DNA fractions were concentrated by ethanol precipitation and digested mildly with Mung bean nuclease in order to remove single stranded DNA tails. The DNAs were then inserted by blunt end ligation into the unique Nru I site of pAT153 plasmid and the two ligation mixture were used to transform E.coli HB101. A total of 67 recombinant clones from (HL) DNA and 158 from (HH) DNA were obtained. The size of the fragments ranged between 200 and 1500 bp with an average of 400 bp.

Analysis of selected clones

From the experiment reported in Fig. 1 we expected to isolate somewhat larger fragments (5000 bp); however considering that Br-containing DNA is

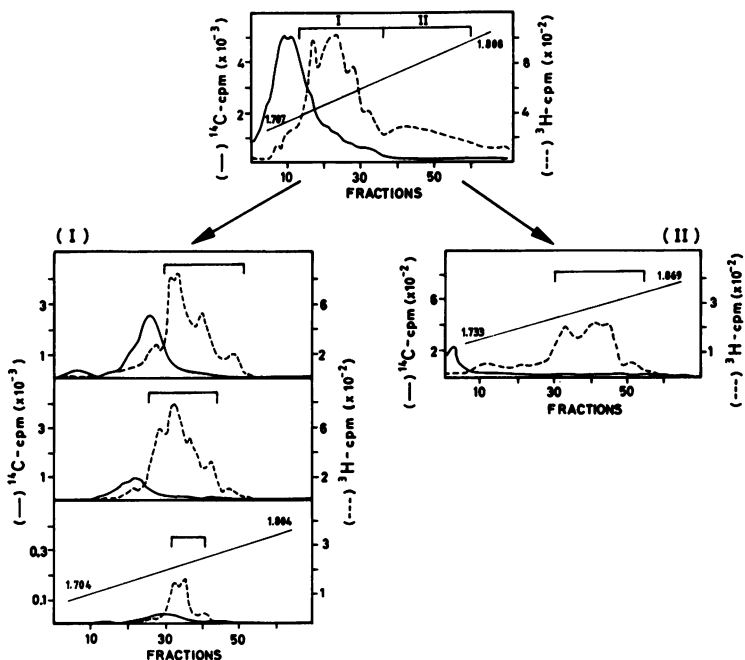


Fig. 3: Separation of newly synthesized DNA by CsCl density gradients. DNA from 4×10^7 cells prepared as described in Materials and Methods was centrifuged in CsCl density gradients ($\rho = 1.75 \text{ g/cm}^3$; $\eta = 1.404$) at 32 000 rpm for 60 hrs at 20°C in a 70 TI Beckman rotor (Quick Seal polyallomer tubes, vol=13ml). 70 fractions of 180 μl were collected from the bottom of the tube and 20 μl aliquots were used for counting the radioactivity and for measuring the refractive index. Fractions corresponding to heavy-light (I) and heavy-heavy (II) DNA were pooled as indicated by brackets. The density gradient is indicated by a straight line whose extremes have the indicated values (g/cm^3).

more easily degraded and the possible side reactions of the treatment with Mung bean nuclease (31) our finding is not surprising. In any case only fragments longer than 400 bp were submitted to further analysis i.e. 15 clones from (HL) DNA and 38 clones from (HH) DNA. In order to assess the frequency of repeated sequences, the (HL) and (HH) DNA clones were assayed by dot blot hybridization with nick-translated Alu, α -satellite DNA, Kpn I and 0-LTR (32-35) probes and with a total nick-translated genomic DNA probe. The results indicated that our selected sample was not significantly enriched in repetitive sequences in agreement with the data on the

renaturation kinetics reported in Fig. 2. In fact only 17 plasmids (10 (HH) and 7 (HL)) showed the presence of repetitive sequences (at least 4000 copies per haploid genomes); of these 4 (3 (HH) and 1 (HL)) contained Alu sequences. None of the 53 plasmids analyzed showed the presence of KpnI, O-LTR or α - satellite DNA sequences. No homology was detected by dot blot hybridization between and within our collections of (HL) and (HH) sequences or between our sequences and the monkey cell ORS sequences isolated with a method similar to ours (16).

Mitochondrial DNA molecules have recently been reported in the preparation of nuclei from several cell lines (36). None of our clones was found to hybridize with human mitochondrial DNA indicating that in our cell line (HL60) and under our conditions, isolation of nuclei prior to DNA extraction efficiently removes mitochondrial DNA contamination.

Verification of the time of replication of selected DNA fragments

The high degree of purification of newly synthesized DNA attained by repeated CsCl density gradient centrifugations should rule out contamination of our sample by randomly assorted chromosomal DNA fragments (LL). However before proceeding to a more detailed characterization, we decided to assess the time of replication of two of the longest fragments (derived from the (HH) DNA collection). Chromosomal DNA replicated at different S-phase intervals (see Materials and Methods) was used in a slot blot hybridization experiment, with nick-translated (HH) and (HL) fragments as probes. For this experiment, only fragments that behaved as single-copy DNA in a Southern blot hybridization to genomic DNA could be used. For this reason the longest (HH) fragment (pB48, 1560 bp) had to be reduced in size (pB48s, see Fig. 5) in order to remove a highly repeated sequence (Alu) present at one end (see below). For the sake of comparison another (HH) DNA fragment from the same sample (pLC46, to be published) was used in the same experiment. As shown in Fig 4 the two fragments (pB48s, 1089 bp and pLC46, 719 bp; lanes 1 and 2) hybridized exclusively to DNA replicated during the first hour (1/3) of the S-phase. The relative signal intensities of these fragments showed that the DNA made in the first hour of the S-phase was highly enriched in these sequences. In the same assay a cDNA fragment of human dihydrofolate

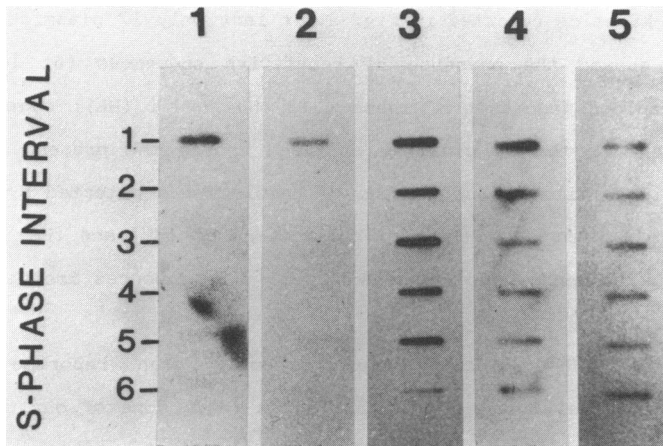


Fig. 4: Time of replication of human DNA fragments.

Preparation of HL60 DNA replicated at different S-phase intervals and slot-blot hybridizations were as described in Materials and Methods. For each S-phase interval 0.1 μ g of (HL) DNA were applied to the slots. The following nick-translated probes were used: lane 1: Pvu II - NruI fragment of pB48s (see Fig. 5); lane 2: a single copy (HH) DNA insert (pLC46, to be published); lane 3: a (HL) DNA fragment than contains a highly repeated sequence (pD8, to be published); lane 4: a 720 bp Pst I fragment of the human DHFR cDNA; Lane 5: a 950 bp Bam HI - EcoRI fragment of the human β -globin gene. The time intervals (1 to 6) are as specified in Materials and Methods.

reductase (DHFR) gene (37) (lane 4) and of human β -globin gene (38) (lane 5) were tested, along with an (HH) DNA fragment containing a highly repeated sequence (pD8 lane 3). As shown in Fig. 4 the DHFR fragment hybridized preferentially to DNA replicated in the first third of the S-phase in agreement with the time of replication of this gene in human cells indicated by other authors (5). Residual hybridization to the later times of the S-phase was probably due to the presence of pseudogenes in the genome. The β -globin gene described as late replicating showed only a limited preferential hybridization at later times (also in this case the presence of homologous genes and of pseudogenes can explain the lack of a clear cut replication time). As expected DNA containing a highly repeated sequence (pD8) strongly hybridized at all times with the exception of the very last one. This result strongly suggests that pB48 and pLC46 inserts actually

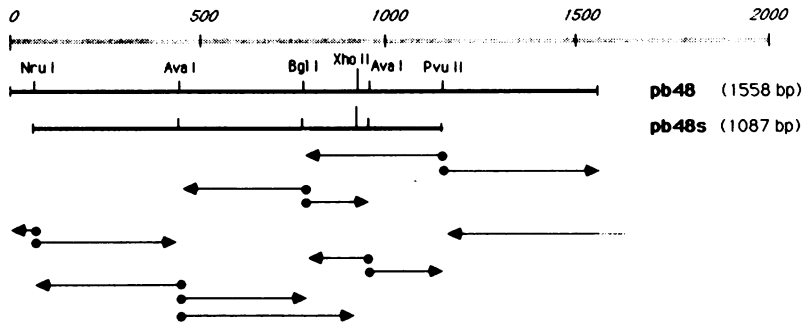


Fig. 5: Restriction map of pB48 human DNA insert and sequencing strategy. pB48s (1089 bp) is a subclone of pB48 (1560 bp) that does not contain the Alu sequence. Dots indicate sites of end-labeling; arrows indicate the direction of sequencing.

derive from the limited subset of replicons activated at the very onset of S-phase. pB48, which was the longest (HH) fragment and gave the strongest hybridization signal to the early replicating DNA, was the object of a more detailed characterization.

Properties of pB48 DNA

The restriction map of pB48 clone is shown in Fig. 5. Initial Southern blot hybridization to genomic DNA had revealed the presence of a highly repetitive sequence. Sequence analysis confirmed in fact that the fragment contained a region homologous to the human Alu sequences (see below) while the remaining part of the insert was single-copy.

By using the sequencing strategy shown in Fig. 5 the complete nucleotide sequence of the pB48 insert (1560 bp) was determined. The sequence (see Fig. 6) had several interesting features that indicate a possible regulatory role. At least three, thermodynamically stable, stem-loop structures can be envisaged (see below and Fig. 7) they that could belong to the snap-back DNA fraction for which the newly synthesized DNA sample is enriched in. The last 189 nucleotides of pB48 (see Fig. 6) correspond to one element (the right half) of the human Alu family (32). In the remaining sequence it is possible to envisage two well defined regions. A GC-rich region (70%), from nucleotides 190 to 750, with the properties typical of the so called "HTF islands" (high frequency of Hpa II sites and of CpG dinucleotides) (39) and

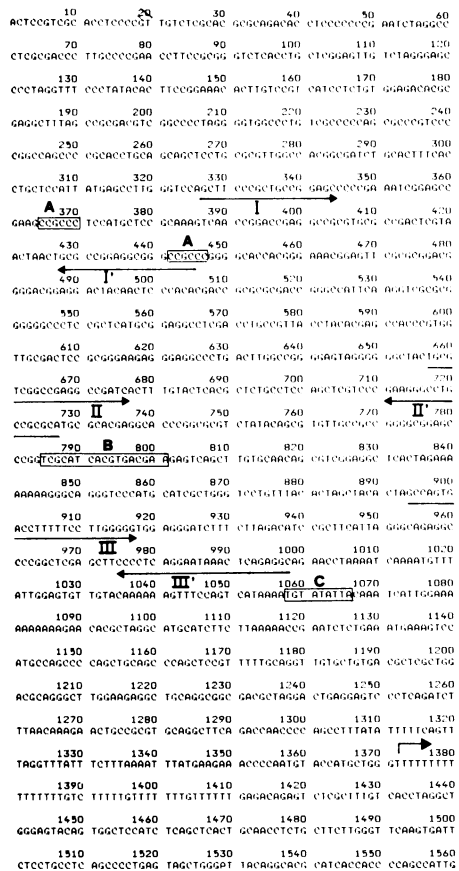


Fig. 6: Nucleotide sequence of pB48.

Major palindromes are indicated by roman numbers (I to III); Boxes: A: Sp1 binding site; B: binding site for the nuclear factor described in the text; C: topoisomerase II binding site consensus. The arrow at nucleotide 1372 indicates the beginning of the Alu sequence.

an AT-rich region (74%), from nucleotides 999 to nucleotide 1092. These two regions encompass a third one (55% GC, nucleotides 751-998) that, as it will be shown in the next section, contains a binding site for a nuclear protein. A second AT-rich region (82%) is present at nucleotides 1304-1351 The pB48 sequence contains only short open reading frames.

A search for direct and inverted repetitions revealed a number of such occurrences. Of the numerous inverted repetitions three, that could give

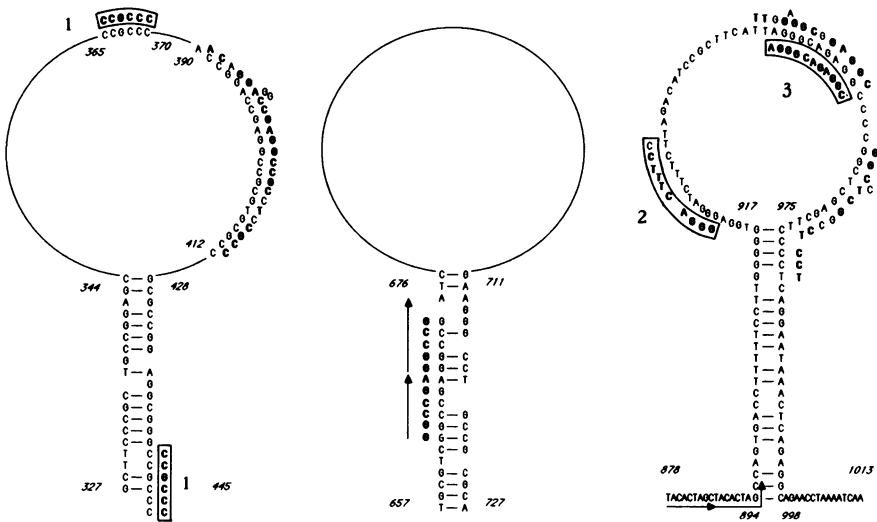


Fig. 7: Thermodynamically stable stem-loop structures in pB48 DNA. The homology with JCV origin of replication on (L) or (U) strand are indicated by bold-typed letters. Arrows indicate directly repeated sequences. Boxed sequences: 1) Sp1 binding site; 2) homology with Ig (k-chain) and SV40 enhancer sequence; 3) homology with a nuclear matrix attachment region of mouse DNA.

rise to relatively stable stem-loop structures, are shown in Fig. 7. In particular, the stem-loop between nucleotides 657-727 contains in the stem a perfect 6 bp tandem repeat, whereas the one between nucleotides 894 and 998 is flanked on the 5'-side by a perfect 9 bp repeat and on the 3'-side by the first AT-rich region. The first stem-loop (nucleotides 327-445) contains two possible Sp1 factor binding sites (40).

Two A-homopolymers (838-845 and 1078-1090) encompass the third stem-loop structure (nucleotides 894-998). Inside the 57 n-loop of this structure, several interesting homologies are observed (see Fig. 7). Nucleotides 922-931 are homologous (70%) to the SV40 and human Ig (K-chain) enhancers (41). Nucleotides 948-979 share 70% homology with the region of the central (GC rich) palindrome in the origin of human papovavirus JCV (42). Two more regions homologous to the same JCV origin sequence are found in the first and the second stem-loop structures shown in Fig. 7; the first (nucleotide 412-390) is 80% homologous while the second (nucleotide 663-673) is 100%

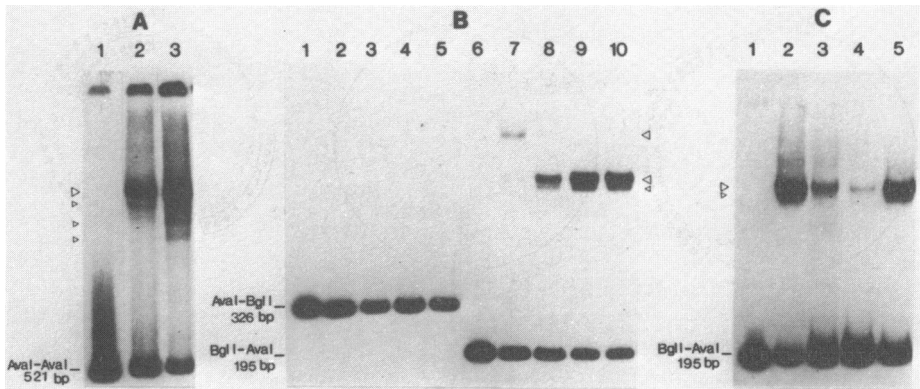


Fig. 8: Localization of the binding site for a nuclear factor in pB48 DNA. All band shift experiments (except those in Panel B) were performed as described in Materials and Methods. Panel A: Band shift assay on the pB48 Ava I - Ava I fragment (521 bp). Lane 1: DNA fragment alone. Lanes 2 and 3: DNA fragment after incubation with HeLa or HL60 cell nuclear extract respectively. Panel B: Band shift assay on the 326 bp Ava I - Bgl I fragment (lanes 1-5) and on the 195 bp Bgl I - Ava I fragment (lanes 6-10) and effect of Mg^{++} . Standard buffer: 20 mM Hepes, pH 7.3, 50 mM NaCl, 0.2 mM EDTA, 5% glycerol. Lanes 1 and 6: without nuclear extract; all other lanes: with nuclear extract. Lanes 2 and 7: standard buffer. Lanes 3 and 8: standard buffer plus 4 mM spermidine. Lanes 4 and 9: standard buffer plus 5 mM $Mg Cl_2$ and 4 mM spermidine. Lanes 5 and 10: standard buffer plus 5 mM $Mg Cl_2$. Panel C: Competition analysis. Lane 1: pB48 Bgl I - Ava I fragment alone. Lanes 2-5: after incubation with HeLa nuclear extract. Lanes 3 and 4 contained also 130 and 520 ng respectively of pB48 plasmid linearized at the Hind III site, corresponding to a 5 and 20 fold molar excess. Lane 5 contained 500 ng of the linearized vector (pAT153). Arrows indicate retarded bands.

homologous. Returning to the third stem-loop structure (nucleotides 894 - 998), nucleotides 950-965 are homologous (83%) to a nuclear matrix associated DNA sequence isolated by Goldberg *et al.* (43) from synchronized mouse 3T3 cells. Such homology raises to 100% when only nucleotides 950-960 are considered. This last region, which shares homology with the binding site III of human BK virus large T antigen, is directly repeated at nucleotides 843-864 (75% homology, rising to 94% homology if the adjacent 8A-homopolymer is included). Lastly, the AT-rich region between nucleotides 999 and 1097 shares extensive similarities with the putative regions of

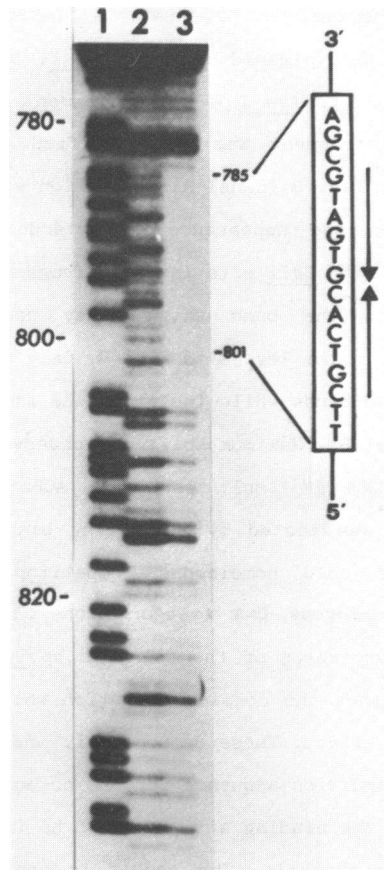


Fig. 9 DNase I footprinting analysis of factor-DNA complex.

The 195 bp BglI - AvaI fragment of pB48 was incubated with HeLa nuclear extract, digested with DNase I and submitted to band shift assay as described in Materials and Methods. The free and the major retarded bands were then eluted from the gel, purified and loaded onto a sequencing gel. Lane 1: G+A sequence ladder. Lane 2: free band. Lane 3: retarded band. The protected sequence is shown alongside.

anchorage to the chromosomal scaffold mediated by the topoisomerase II as described by Laemmli (44).

Band shift assay in gel electrophoresis and DNase I footprinting

In view of the indications that pB48 insert could contain possible regulatory signals we searched for binding sites for specific nuclear

factors by employing the band shift assay. A preliminary experiment was performed on the entire pB48 plasmid cut with Ava I. Of the three fragments so obtained, only the 521 bp Ava I - Ava I fragment seemed to be retarded (data not shown). This fragment was then purified and used for further experiments. As shown in Fig. 8 (panel A) incubation with both HeLa and HL60 nuclear extracts causes the appearance of retarded bands. The 521 bp fragment was then cut at the BglI site into two fragments of 326 and 195 bp that were submitted to the band shift assay under different buffer conditions as specified in the legend to Fig. 8. Only the 195 bp BglI - Ava I fragment shows retarded bands while the other one is unaffected under any conditions (Fig. 8, panel B). Maximum shift was observed in the presence of 5 mM Mg^{++} . Spermidine (4 mM) only partially substitutes for Mg^{++} . The specificity of binding was tested by performing binding reactions in the presence of excess of cold homologous competitor DNA (plasmid pB48, linearized) and non homologous DNA (vector alone). As shown in Fig. 8, (panel C) a 20 fold molar excess of the cold 195 bp Bgl I - Ava I fragment almost completely abolishes the complex formation while the same excess of heterologous DNA has no effect. These data clearly indicate that the 195 bp fragment contains a recognition sequence for one or more nuclear factors.

In order to localize the binding site, the 195 bp fragment was submitted to DNase I footprinting analysis. The result is shown in Fig. 9 where a protected region of 17 nucleotides is visible on the L-strand between nucleotides 801 and 785. Interestingly the protected sequence: 5'TTCGTCACGTGATGCGA3' is palindromic in the 12 central nucleotides (except for one mismatch) and shows a remarkable homology with an upstream element in the major late promoter (MLP) of Adenovirus type 2 (26).

DISCUSSION

Kinetics of early DNA replication

The data reported in Figure 1 indicate that under our experimental conditions new DNA chains are initiated which elongate at a rate of about 2 base per second per fragment (one base per second at each end, assuming bidirectional replication). By comparison with the normal rate of fork

advancement (50 bases per second) we estimate in our conditions an inhibition of the rate of elongation down to 2% residual, in agreement with the known inhibition of DNA polymerase and of growing fork advancement by 1 µg/ml aphidicolin. As shown in Table 1, we can estimate the number of newly synthesized fragments per cell at different times following release from aphidicolin block. Such a number is roughly constant from 10 to 130 min (approximately 1500 fragments per cell). It is therefore conceivable to assume that about 750 replicons are activated per cell and that their activation occurs in the first minutes.

Properties of the newly synthesized DNA

Two features appear prominent in the DNA synthesized at the onset of the S-phase: 1) it is significantly enriched in snap-back sequences (see Fig.2); 2) a substantial amount of (HH) DNA is observable at very early times (see Fig.3). The presence of (HH) DNA can be explained either by extrusion of nascent DNA molecules (28) or by reinitiation after prolonged inhibition of DNA replication (23). The analysis of the cloned fragments did not reveal extensive homologies or enrichment for highly repeated sequences. It must be remembered however that relatively short repeated sequences, or longer partially divergent ones would have escaped detection by our methods.

Properties of pB48 fragment

The longest early replicating DNA fragment (pB48) shows a number of distinctive features that point to its having regulatory roles in the cell. The sequence contains three palindromes, with stability greater than -27 Kcal/mol. Thus, about 8% of the sequence can be considered operationally "snap-back DNA". This is to be compared with 3.5% of the overall DNA, and 5.6% of the DNA synthesized within the first 7 minutes of S phase. The three possible stem-loop structures (Fig.7) contain regions of homology to the central GC-rich palindrome in the origin of human Papovavirus JCV, to sequences associated to the nuclear matrix in synchronized mouse cells and to the SV40 and human Ig(k-chain) enhancer (see Fig. 7). Between the second and the third stem-loop lies a binding site for a nuclear factor, homologous to an upstream element in the adenovirus major late promoter. Our sequence could be one of the cellular targets of the factor (USF), described to bind

such an element and to stimulate transcription (26). Finally, outside the palindrome-rich area, an AT-rich region is found that contains a topoisomerase II consensus binding site, as well as other features typical of scaffold attachment regions (SAR) in animal DNA (44). There are several indications that SAR sequences might have an important regulatory function in the activation of transcription and of DNA replication.

In conclusion our results support the hypothesis of a tight interdependence between initiation of replication, transcription and binding to nuclear structures (4). In this respect it is important to notice that preliminary studies on another early replicating fragment (pLC46, see Fig.4) also evidenced the presence of transcriptional signals and SAR regions (to be published). The demonstration of the biological role of our sequences requires however a functional assay in vivo. Experiments in this direction carried out in our and in other laboratories have so far given negative results. However, considering some recent reports (45, 46) on the isolation of DNA segments that can autonomously replicate in mouse and human cells we will further pursue the demonstration of an in vivo activity of our sequences.

AKNOWLEDGMENTS

Research carried out in the framework of Contract N° BAP-0222-I of the Biotechnology Action Program of the Commission of European Communities and supported by the "Progetto Finalizzato Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie" and by the "Progetto Finalizzato Oncologia", C.N.R., Rome. The authors wish to thank Dr. Fabio Cobianchi and Dr. Guido Pedrali-Noy for assistance in some experiments, Dr. R.G. Martin (N.I.H.) for providing unpublished information and for helpful discussions and Mrs Cristina Bagarotti for typing the manuscript.

*Present address: International Centre for Genetic Engineering and Biotechnology, Padriciano 99, I-34012 Trieste, Italy

REFERENCES

- 1) Taylor, J.H. (1960) *J. Biophys. Biochem. Cytol.* 7: 455-464.
- 2) Huberman, J.A. and Riggs, A.D. (1968) *J. Mol. Biol.* 32: 327-348.
- 3) Blumenthal, A.B, Kriegstein, H.J. and Hogness, D.S. (1974) *Cold Spring Harbor Symp. Quant. Biol.* 38: 205-223.
- 4) Harland, R. (1981) *Trends Biochem. Sci.* 6: 71-74

- 5) Goldman, M.A., Holmquist, G.P., Gray, M.C., Caston, L.A. and Nag, A. (1984) *Science* 224: 686-692.
- 6) Palen, T.E. and Cech T.R. (1984) *Cell* 36: 933-942.
- 7) Vogelstein, B., Pardoll, D.M. and Coffey, D.S. (1980) *Cell*. 22: 79-85.
- 8) Smith, H.C., Puvion, E., Buchholtz, L.A. and Berezney, R. (1984) *J. Cell Biol.* 99: 1794-1802.
- 9) Jackson, D.A. and Cook, P.R. (1986) *The EMBO J.* 5: 1403-1410.
- 10) Wold, M.S., Li, J.J. and Kelly J.T. (1987) *Proc. Natl. Acad. Sci. USA* 84: 3643-3647.
- 11) Watanabe, N., Akasaka, K., Shiroya, T., Obinata, M. and Shimada, H. (1986) *Nucleic Acids Res.* 14: 9509-9515.
- 12) Kornberg, A. (1980) *DNA replication*. Freeman and C. San Francisco USA.
- 13) de Villiers, J., Schaffner, W., Tyndall, C., Lupton S. and Kamen R. (1984) *Nature* 312: 242-246.
- 14) Lusky, M. and Botchan, M.R., (1986) *Proc. Natl. Acad. Sci. USA* 83: 3609-3613.
- 15) Biamonti G., Della Valle, G., Talarico, D., Cobianchi, F., Riva, S. and Falaschi, A. (1985) *Nucleic Acids Res.* 13: 5545-5560.
- 16) Kaufmann, G., Zannis-Hadjopoulos, M., and Martin R.G. (1985) *Mol. Cell. Biol.* 5: 721-727.
- 17) Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning*. Cold Spring Harbor Laboratory, N.Y. USA.
- 18) Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* 132:6-13.
- 19) Pedrali-Noy, G., Spadari, S., Miller-Faures, A., Miller, A.O.A., Kruppa, J. and Kock, G. (1980) *Nucleic Acids Res.* 8: 377-394.
- 20) Challberg M.D. and Kelly T.J. Jr. (1979) *Proc. Natl. Acad. Sci. USA*, 76, 655-659.
- 21) Hanahan, D. (1983) *J. Mol. Biol.* 166: 557-580.
- 22) Pedrali-Noy G. and Weissbach A. (1977) *J. Virol.* 24: 406-407.
- 23) Mariani, B.D. and Schimke, R.T. (1984) *J. Biol. Chem.* 259: 1901-1910.
- 24) Maxam, A.M. and Gilbert, W. (1980) *Methods in enzymology* 65: 499-560.
- 25) Dignam, J.D., Lebovitz, R.M., and, Roeder, R.G. (1983) *Nucleic Acids Res.* 11: 1475-1491.
- 26) Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) *Cell* 43: 439-448.
- 27) Studier F.W. (1965) *J. Mol. Biol.* 11: 373-390.
- 28) Zannis-Hadjopoulos, M., Kaufmann, G. and Martin R.G. (1984) *J. Mol. Biol.* 179: 577-586.
- 29) Schimke, R.T., Sherwood S.W., Hill, B.A. and Johnston R.N. (1986) *Proc. Natl. Acad. Sci. USA* 83: 2157-2161
- 30) Zannis-Hadjopoulos, M., Persico, M. and Martin R.G. (1981) *Cell*. 27: 155-163.
- 31) Kowalski, D., Kroeker, W.D. and Laskowski, M. (1976) *Biochemistry* 15: 4457-4467.
- 32) Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T. and Schmid, C.W. (1981) *J. Mol. Biol.* 151: 17-33.
- 33) Shmookler Reis, R.J., Srivastava, A., Beranek, D.T. and Goldstein, S. (1985) *J. Mol. Biol.* 186: 31-41.
- 34) Grimaldi, G., Skowronski, J. and Singer, M.F. (1984) *The EMBO J.* 3: 1753-1759.

Nucleic Acids Research

- 35) Paulson, K.E., Deka, N., Schmid, C.W., Misra, R., Schindler, C.W., Rush, M.G., Kadyk, L. and Leinwand L. (1985) *Nature* 316: 359-361.
- 36) Kristensen T. and Prydz H. (1986) *Nucleic Acids Res.* 14: 2597-2609.
- 37) Morandi, C., Masters, J.N., Mottes, M. and Attardi, G. (1982) *J. Mol. Biol.* 156: 583-607.
- 38) Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1979) *Nature* 279: 598-603.
- 39) Bird, A.P. (1986) *Nature* 321: 209-219.
- 40) Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) *Trends Biochem.Sci.* 11: 20-23.
- 41) Sen, R. and Baltimore D. (1986) *Cell.* 46: 705-716.
- 42) Frisque, R.J., Bream, G.L. and Cannella M.T. (1984) *J. Virol.* 51: 458-469.
- 43) Goldberg, G.I., Collier, I. and Cassel, A., (1983) *Proc. Natl. Acad. Sci. USA* 80: 6887-6891.
- 44) Gasser, S.M. and Laemmli U.K. (1987) *Trends Genet.* 3: 16-22.
- 45) Ariga, H., Itani, T. and Iguchi-Ariga, S.M.M. (1987) *Mol. Cell. Biol.* 7: 1-6.
- 46) Iguchi-Ariga, S.M.M., Itani, T., Kiji, Y. and Ariga, H. (1987) *The EMBO J.* 6:2365-2371.